

UNIVERSITY

College of Humanities
& Social SciencesDepartment of Psychology
& Sociology
(215) 895-2455

MAY 2 - 1976

IEEE - EDITORIAL
NEW YORK

April 26, 1976

Mr. Robert W. Lucky, Editor
 Proceedings of the IEEE
 345 East 47th Street
 New York, New York 10017

Dear Mr. Lucky:

I have just read, with considerable interest, an article in the March issue of the Proceedings of the IEEE by H. E. Puthoff and R. Targ entitled "A Perceptual Channel for Information Transfer Over Kilometer Distances: Historical Perspective and Recent Research." These authors have undertaken a series of potentially significant studies into a topic of obvious interest to an increasingly large number of individuals, doing so utilizing what appears to be appropriate techniques taken from the scientist's arsenal of objective methodologies, many of which the authors have previously distinguished themselves using in research in the natural sciences. It is certainly an understatement to say that such a controversial topic as ESP requires extraordinarily precise and careful methods, especially well suited to exclude the basis for all those nagging criticisms often irreverently hurled at the parapsychologist--fraud, inadequate controls, imprecise and incomplete reporting of details, improper statistics, proper statistics improperly used, and, especially, the presence of numerous confounding variables which vary with the independent variables in their experimental designs in such a way that these other uncontrolled variables remain to potentially account for the given results. Indeed, the authors themselves apparently set out to achieve this goal, what they refer to as their "principal responsibility--to resolve under unambiguous conditions the basic issue of whether or not this class of paranormal perception phenomenon exists" (pp. 334-335). Unfortunately, the model for their study follows the traditional stratagems of the parapsychologists in the United States and Europe, rather than the method they no doubt otherwise use in their non-behavioral research in the natural sciences (and I might add that is largely followed in experimental psychology). As we shall indicate, the consequence of this is that they must necessarily fall far short in fulfilling their stated "principal responsibility."

It is the essence of the experimental method--in contrast to naturalistic, observation, the survey technique, correlational procedures, field studies, and the "Theoretical Model" of the parapsychologists (Girden, 1962, p. 360)--to in fact create



the conditions necessary for "unambiguous" resolution of fundamental questions since only this method permits manipulation and control of potentially confounding variables by the eminently sensible method of varying the critical factors under study, and systematically observing their effects upon other selected and measurable variables while holding these potentially confounding variables under tight control via such techniques as randomization, constancy, counterbalancing, matching, etc. The manipulated variables are called independent variables (IVs), and the variables sensitive to the effects of the IV is called the dependent variables (DVs). This has been expressed quite eloquently by Ebbinghaus (1964, p. 7; original: 1885) in his incorporation of this experimental method into his researches into human memory, a successful effort which went far in illustrating in 1885 the power of this method in scientifically understanding human behavior--including perception:

"We all know of what this method consists: an attempt is made to keep constant the mass of conditions which have proven themselves causally connected with a certain result; one of these conditions is isolated from the rest and varied in a way that can be numerically described; then the accompanying change on the side of the effect is ascertained by measurement or computation."

The simplest experiments, therefore, are those employing but a single IV and a single DV, and, in the fundamental situation in which an effort is being made to demonstrate the sheer existence of a phenomenon (as in the present study, without inquiring further into its composition and contingencies), the two basic values or levels or variates of the IV may be simply designated the "experimental" (i.e., the factor appears--operationally defined--in some amount) and "control" condition (i.e., the factor appears as a zero amount).

A helpful example might be a drug study in which we only wanted to know if the drug makes running a maze difficult--or not! In this situation, we would experimentally compare the drug condition's effects to a condition otherwise identical in which, however, the drug is absent! This would constitute one of the most fundamental empirical control conditions used in experimental science in general, and experimental psychology in particular, although of course other empirical control conditions are possible, depending on what one is controlling for. We speak, by the way, of "control condition", and not "controlled" condition, since we assume all conditions in the experiment are controlled in some way. This is apparently a source of much confusion in the parapsychological literature, where constant reference is made to "controlled laboratory conditions," "scientifically controlled conditions," and so on, thereby creating the illusion that basically well conceived control conditions are being used. In the present example, if one is concerned that the procedure of injection causes errors to be made on the maze, then we need a placebo control in which S is treated identically as in the drug condition, but saline rather than the drug is injected. In this case, we "control for" the injection procedure's effects on the DV by holding this potentially confounding variable constant and can thereby evaluate it and prevent it from creating the erroneous impression that the drug, per se, produced the errors S made on the maze, when in fact the drug injection procedure itself may have produced that effect on the DV.

As we know, in the present study, the basic procedure described was carried out with all the subjects (six in number for section III studies through subsection "D",

two more in section "E", and five more in section "F"; we limit our critique to the more detailed accounts given in sections A-D rather than the very sketchy material in sections E and F, although all these studies used essentially the same procedure and varies mainly on the basis of subject characteristics.). That is, all the subjects were administered, as it were, the same basic treatment condition, and were thus all part of the same "clairvoyant" or "remote viewing" group, consisting essentially in the Ss making an effort to somehow envision a remote target. Actually, all we know of the instructions to the S, so critical in determining the operational definition of this condition, is that the "remote-viewing subject was asked to describe his impressions of the target site into a tape recorder and to make any drawings he thought appropriate (p.335)," since the authors do not give us the actual instructions. In any case, for each S, there did indeed exist a "remote target" designated by a "target team" or "demarcation team" situated at some geographical location nearby the SRI laboratory, although the theoretical rationale for this team's existence and its role in the procedure is never really logically presented by the author, save only for an historical precedent established working informally with Mr. Ingo Swann (p. 334) and perhaps the anecdotal "pilot experiment" (p.330) where one of the authors functioned in this capacity as a "demarcation team" and, we infer, grew quite excited and impressed with a subject's descriptions of some sites he visited, where the subject supposedly had no particular prior information about the site in question. This situation, we shall detail later, results in considerable ambiguity in the designation of the "target" (is it the "target-perceived-by-the-team", or is it the physical stimulus of the target, and so on?).

The whole procedure, carried out under what the authors describe as "rigidly controlled scientific conditions (p.334)", and including the subsequent judging procedure we shall discuss in some detail later, could have been called an "experimental" condition comparable to the "drug present" condition in the example above if there had been a control condition with which to compare it (comparable to the "placebo" or "no drug" conditions in our example), that is, a condition in which all of the preceding procedures were exactly followed but in the absence of what was previously operationally and objectively defined as the "remote-viewing" condition (note the importance of objective specification of the conditions under which the "remote-viewing" is to occur, because without that there can be no objective controlled variation of the condition since one would never know quite when the experimental condition existed in the first place!). As the matter stands, since only one condition was run in this study, we really have neither experimental nor control conditions (because the terms are defined relative to each other); and since that is the case, we really have no IV; and since we have no IV, we actually do not even have one experiment (let alone the 50 or so claimed by the authors in this single publication--p. 330).

The DV, of course, also deserves careful scrutiny, since hypothesized IV effects are evaluated in terms of changes in DV measures. In this study, the actual "number" obtained as the basic raw datum was the individual ranking or "match" number of a subject's tape recorded description (supposedly of a "target"), with some aspect (perhaps) of a nearby geographical location. The number could assume any value from "1" to "9", where "1" referred to the judge's estimate of a "best" match, through "9" which was assigned to a match if it was a "worst" match for a given target. Nine targets and nine descriptions were obtained for each S. This was apparently an ordinal scale of measurement. We must assume, in the absence (also) of a specific report of the instructions to the judge, that the judge knew exactly what his task was in determining the degree of correspondence between the Ss' descriptions and the judge's own unspecified perceptions when physically present at the so-called target.

We should, therefore, consider the precise meaning of a "hit" under these circumstances. Evidently, a "hit" for the Es occurred when the geographical locus they called the "target" is judged to correspond to the S's description obtained while the "demarcation team" was visiting a particular site. This was also, therefore, a situation where the S's corresponding description was assigned a rank of "1". Since 9 targets were "experimented" with a given S, we have 9 separate judgments of rank carried out by a single given judge. The sum of ranks for the given S of all the descriptions for the theoretically associated targets was then compared to a probability distribution for that statistic, a situation we claim is really improperly serving as a substitution for a test based on some statistic (which could be the sum of ranks) obtained from the comparison of the matches for this hypothetical "remote viewing" data to the matches obtained under similar judging conditions where there was no remote viewing possible (other control conditions will be described below).

However, there are numerous peculiarities about this judging process possibly unresolvable on the basis of the authors' scanty and ambiguous description of this critically important aspect of their study. Let us consider a few of these problems.

1. What was the judge judging? The judge for a given S's performance for a given target (what the authors refer to incredibly as an "experiment") was successively driven to each geographical location previously visited by the peripatetic Es. Since, as we previously noted, we do not know precisely what aspects of the geographical location constituted a "target" in the original "experiment" when the demarcation team was present, and since it is even more ambiguous now what the judge was viewing, as well as what he was supposed to be looking at while he reviewed the S's packages of 9 descriptions, we seem in this procedure, therefore, to actually be dealing with two (and perhaps three) recognizably distinct categories of "targets": one is constituted by the perceptions of the demarcation team; a second by the perceptions by the judges; and a third by direct physical aspects of some geographical location (photographs are used in the report and labelled "target" to further complicate target delineation--e.g., Figure 4). It is difficult to evaluate how potentially dissimilar these various "targets" were in the absence of clarifying and detailed accounts of the specific instructions to the teams and to the judge. After all, the judge or demarcation team may have fixated the horizon, focused on passing vehicles, noticed a sign, and so on. Consequently, the reporting of the target simply as a "target location", or "remote location" (e.g., "Marina, Redwood City") as the authors do, belie the fact that we are dealing here with both their clairvoyance belief and in the perceptual theory that the way things appear to a viewer and the way they are physically constituted may not be at all identical (e.g., the sensation "red" of a perceived stop sign, and the atomic structure of the paint pigments on its metal surface). The possible consequence of this lack of precise definition of the target is that the judge has considerable latitude in fitting S's description to the scene as he perceives it, possibly even looking for certain aspects of the scene before him that seem to be present in some of S's descriptions.

2. How reliable is the Judge's "Matching" Measure? A direct consequence of the preceding consideration bears directly upon the integrity of the measurement used for the authors' basic datum, namely, the problem that we do not know specifically what constitutes the "standard" (the target" in some sense) against which our "comparison stimuli" (S's descriptions) are measured by the judge, resulting in what must be an impressively unreliable measure (DV). Furthermore, this potentially highly unreliable

measure, therefore, becomes potentially even more unreliable when one recalls that only one judge was used for each subject's data, and he performed only one judging (a series of nine evaluations) for a given target. The report tells us nothing of any efforts to "train" the judge in this procedure, or otherwise to determine some initial reliability coefficient for this kind of task for a naive judge. This question of doubt about the status of our measuring instrument, our "sensor" as it were, is especially perplexing in light of the highly refined and instrumentally sophisticated dependent variables the authors surely use in all their other research in the physical sciences; why they have avoided use of such techniques in this so-called "remote-viewing" research, and furthermore not used the appropriate control procedures experimental psychologists and psychophysicists have laboriously developed starting with Wundt's work in Leipzig in 1879, remains rather enigmatic, to say the least.

3. Were the Judgments Independent? After the first "target" site was visited, we know the judge was taken, according to some unspecified procedure (the authors say "in turn," but "in turn" of what--demarcation team visits, random sequence just for the purposes of judging, or simply (and necessarily) "one after the other?") to another target. Was he then given the same nine descriptions in precisely the same way? If so, how was the procedure handled? Were precautions taken to prevent cues associated with previous correct 'hits' from subsequently influencing later matches, since, obviously, information that a previous description had been correctly matched (a "hit") reduces by one those descriptions in the pool that he draws from to "match," thereby significantly increasing the probability of making subsequent hits. This could easily occur if an informed E drove the judge to the different sites for the judging, and then, having noticed what we would assume to be an accidental "hit" by the judge, indicated as much through perhaps an unconscious cue to the judge. The point is, not that we think the authors really overlooked such an obvious source of bias, but rather that we need special reassurance specifically that a naive driver with explicit driving instructions was used to merely drive the judge to the different locations--and that is all! After all, he cannot cue the judge about previous accomplishments, if he knows nothing of the research. Of course, the judge may have driven himself to the site, which would largely eliminate this source of confounding in the measurement series; but we do not know what happened here since these pertinent details are omitted from the report.

To summarize our concerns with the dependent variable to this point, we may say, first, it appears to need considerable specification, a detailed operational definition; second, its reliability should be ascertained, and doubt that it is inherently very unreliable removed; third, it should be used for determining experimental condition ("Clairvoyance") hit rates in the matching task, as well as for appropriate control conditions! ("No clairvoyance," "no clairvoyance and no demarcation team," and so on) hit rates, so that an appropriate statistic based on the net differences in "hit rates" could be tested for the degree to which this difference measure is statistically significant.

Since this idea of an appropriate control condition seems to be so alien to the parapsychological so-called "experimental" literature, but is everywhere else in experimental science recognized to be of the essence of the logic of experimentation in the first place, let us consider in some detail what the minimal requirements would be in order to actually run an experiment on this "remote viewing" hypothesis, rather than merely report curious observations of possibly purely coincidental events.

Apparently, "clairvoyance" seems to be viewed or interpreted as some kind of non-sensory "seeing clearly" of an object not immediately available to sensory surveillance and scanning. It would seem, therefore, that for objective experimental purposes it may be operationally defined either by specifying instructions to a subject concerning his potential task of envisioning distant objects, or by controlling the very existence of a target to be seen "remotely" (or any other way for that matter). Therefore, each of these procedural details can serve as an IV in a well controlled 2-factor design. For the "target" IV, we would seem to need only the two basic levels of this IV, namely, a condition in which a remote target is available for alleged "remote viewing," hypothetically facilitated in some inscrutable fashion by a "demarcation team" actively observing the target according to explicit, replicable directions, and a condition in which no such remote target or team exists. The former instance would theoretically permit clairvoyance to operate (if it exists), since one's reported imagery could be matched or compared to the actual selected target, as discussed in detail above; the latter control condition would not, since there would be no selected target to see clairvoyantly or otherwise.

Furthermore, on the basis of an operational definition of remote-viewing in terms of the instructions to S establishing his remote-viewing task, we can proceed to require him to draw pictures or verbally describe anything he can imagine for, say, a 5-minute period (commencing and ending at our signal), but in the experimental condition further telling him to attempt to view a remote target, whereas refraining from such a demand in the control condition. The overall design then would be:

INSERT DIAGRAM HERE

Procedurally, this would mean we begin by creating a large pool of nearby geographical sites to function as potential "targets" (This is an attempt to preserve the authors procedures wherever feasible, since it would hardly otherwise be very advisable to use such complex stimuli as these targets when little is known of these effects with simple stimuli), and then randomly assign targets to the 4 cells of our experimental design until, say, each cell has five (5) targets (or perhaps more depending somewhat on resources and the type of statistical test to be used). As for our "instruction" variable, we will basically require S to give a detailed phenomenological report of his "stream of consciousness" experiences at designated moments, but, for half of the targets (cells A and B) we would have in addition asked S to attempt to "visualize" some geographical scene (we would provide our readers in the later report with detailed, replicable, instructions so the reader could evaluate the degree to which we may have given our S any information about the type, condition or location of the target, all of which information may help diminish the population of total targets S thinks might be under consideration, thus significantly increasing the probability of making a hit--a circumstance that may help account for such outstanding "hits" in the authors' report as that of White Plaza). The differential

TARGET

IV #2

		PRESENT (EXPERIMENTAL)		ABSENT (CONTROL)	
		B	A	C	D

REMOTE-VIEW
(EXPERIMENTAL)

DO NOT
REMOTE-VIEW
(CONTROL)

DV: NUMBER OF CORRECT
MATCHES ("HITS")

presence of such information, varying with the "instruction" IV, would be an excellent example of a confounding variable in an experiment of this type. In any case, the use of these control conditions (no target, no remote-viewing demand in various combinations designated by cells A, C and D) permits an empirical determination of the extent to which such "hits" in matching may occur simply by virtue of the similarities judges may read into this matching material due to its inherent unstructured subjective quality, much as human observers of Rorschach ink blots see a map of Ireland or a bust of Caesar, an organization imposed on perceptual materials that perceptual theorists refer to as "Gestalts," often reflecting unconscious needs of the observer (such as, "I have psychic powers I must demonstrate!").

As for the sequence to be followed in actually running S through these different treatment combinations, a simple randomized arrangement would suffice, that is, assuming 20 targets, with numbers 1-5 in cell A, 6-10 in cell B and so on, we simply run the treatment combination according to the next number in the random series (e.g., if 13 is the first random number, we run S according to the conditions obtaining for cell C). Later the judge can in random sequence consider one cell block of descriptions and targets after the other, until all four cells have been "matched," where naturally the judge is not informed of the treatment combination condition. Hit rates are determined, for example, and then appropriate statistical tests for simple effects, main effects, and possible interactions determined. The "remote-viewing" hypothesis would be confirmed if cell "B" had significantly more hits than cell "A" (no target), cell "C" (target, no R-V instruction), and certainly than cell "D" (where both target and appropriate instructions are absent). The present reported study, by the way, consists only in cell "B", making appropriate comparison impossible. In any case, lack of such significant differences will merely disconfirm this parapsychological hypothesis (not "disprove"), thereby providing no one with any empirical reason for affirming belief in the existence in this unusual phenomenon's existence--at least, as evaluated in this suggested experimental design.

Furthermore, the underlying logic should be clear, that the clairvoyant hypothesis cannot be evaluated when matches or hits (cell "A") are compared to some hypothetical "chance" level (no matter how accurate the statistical test be according to individual statisticians, pronouncements of statistical associations, or as found in statistics texts) that fails to represent the normal or average number of hits that might be really expected under these identical conditions when clairvoyance does not exist (as cells "D" especially, and also "A" and "C" establishes, as explained above).

The essential experimental invalidity of this parapsychological tactic of comparing their so-called "experimental data" to some hypothetical "chance" level, especially in the absence of experimental control conditions, is perhaps most clearly illustrated in the psychokinetic ("mind over matter") literature. In those studies, classically championed by Rhine some years ago (e.g., 1947), a tumbler might toss out onto a table some 600 dice; the S being evaluated for his alleged "psychic" skills makes some effort to mentally or "psychically" influence each die and have, say, as many 5s turn up as he can. A tabulation is made, and it is found that, say, 235 5s have indeed turned up--a startling result, indeed, especially when it is considered that only 100 should have so turned up by "chance." The problem of this interpretation is the same as that in the present "remote-viewing" study. This "chance level" is a theoretical, mathematical abstract model of the behavior of 600 ideal dice, not

necessarily related to the die actually being used in the study at all, not to mention to the procedures and physical conditions operating in actually using and "tossing" them. An empirical control condition is necessary to obtain this "standard" for comparison, one representing the actual concrete die used in the study, and accounting for such potentially confounding variables influencing the throw such as non-horizontal tables, loaded dice, etc. (all as discussed above). This control condition would exactly duplicate the experimental condition, save only for the omission of some critical factor under study in the experiment, such as (in this case) psychokinesis itself! In other words, we run a non-PK condition, in which all is done as in the experimental condition, but S does not "PK" when the die is thrown.

Would not such a comparison change our interpretation of this whole situation if it was found that again about 230 dice came up "5s"? We would then obviously look for those other alternative, non-parapsychological explanations for the high 5's count (such as the possibility of loaded dice). An excellent review of the PK literature through 1962, and general critique of this "theoretical model", is that of Girden (1963).

- It may be true, as the parapsychologists claim, that ESP is real and represents a great latent power of the human mind, one day to emerge in full recognition by science as another momentous step in the evolution of man and his mind; but its truth remains to be demonstrated through use of the experimental method, and until it is, in the same way as Ohm's law or Pavlov's conditioned reflex paradigm has been so demonstrated, parapsychologists ought not consider psychologists and other scientists and engineers calcified conservatives blindly refusing to see the obvious "fact" that ESP, etc., exists, because the parapsychologists themselves seem to virtually intentionally avoid using the only techniques which in the long run will prove persuasive to the scientific community, and those are objective experimental procedures.

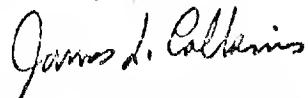
Perseverance of belief in ESP and related phenomena built upon such shabby and preposterous evidence as is usually offered as "scientific proof" on behalf of the ESP proponents (and the present study is to some extent an example) itself requires some explanation, since such perseverance is often cited as itself somehow evidence for the existence of these "paranormal" phenomena (important people wouldn't believe it if it wasn't true!). It is not hard to find such an explanation in the literature of the social psychology of social movements and cults, since therein it is well understood how organized groups of individuals may band together with their own ideologies, their own clubs, their own in-group publications and sacred works, their own symbols, passwords and slogans, in the interest often of providing some sense to life, some direction, some compensation for a sense of personal loss, insignificance, or inferiority, which is provided by becoming a "true believer," as Hoffer (1951) put it, in some special and unique movement. Especially is this true of ideologies that persist despite virtually universal rejection on some rational grounds, suggesting, in contrast, that the belief in question answers to needs other than those that are rational. The particular difficulty the parapsychologists seem to encounter in that (most fortunately for humanity) the scientific method represents one of the few really pristine exemplars of the rational use of the human mind, thereby guaranteeing a clash with their own vested irrational ideological systems. Add to this the consideration that we are presently witnessing a strong resurgence of popular interest in the occult and supernatural, with a proportionate increase in the volume of superstititious behavior prevalent, we can see that the

present study decidedly fits into the "Zeitgeist" of contemporary dementia. This is especially unfortunate for psychology, since this also means considerable talent and potential expertise that could have pushed back frontiers of new understanding of the normal complex realities of human perception and its relation to the nervous system, and physical reality, should instead be diverted into inquiry into the bizarre circuitous vagaries of the so-called "para-normal".

REFERENCES

1. Ebbinghaus, H. Memory: A Contribution to Experimental Psychology. New York: Dover Publications, 1964 (original: 1885)
2. Girden, E. A Review of Psychokinesis (PK). Psychological Bulletin, 1962, 59, 353-388.
3. Hoffer, E. The True Believer. New York: New American Library, 1964 (original 1951).
4. Puthoff, H.E., and Targ, R. A Perceptual Channel for Information Transfer over Kilometer Distances: Historical Perspective and Recent Research. Proceedings of the IEEE, 1976, 64, 329-354.
5. Rhine, J. B. The Reach of the Mind. New York: W. Sloane Associates, 1966

Sincerely yours,



James L. Calkins, Ph.D.
Associate Professor

JLC:rm